



Porter, M. L., Atawa, H., Bok, M. J., & Cronin, T. W. (2020). Exceptional diversity of opsin expression patterns in *Neogonodactylus oerstedii* (Stomatopoda) retinas. *Proceedings of the National Academy of Sciences of the United States of America*.
<https://doi.org/10.1073/pnas.1917303117>

Peer reviewed version

Link to published version (if available):
[10.1073/pnas.1917303117](https://doi.org/10.1073/pnas.1917303117)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via National Academy of Sciences at <https://www.pnas.org/content/early/2020/04/01/1917303117> . Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

Exceptional Diversity of Opsin Expression Patterns in *Neogonodactylus oerstedii* (Stomatopoda) Retinas

Megan L. Porter^{*1,2}, Hiroko Awata^{*1}, Michael J. Bok^{1,3}, Thomas W. Cronin¹

1 – Department of Biological Sciences, University of Maryland Baltimore County,
Baltimore, MD 21250

2 – current address: Department of Biology, University of Hawai'i at Mānoa, Honolulu,
HI 96822

3 – School of Biological Sciences, University of Bristol, Life Sciences Building, Bristol,
BS8 1TQ, UK.

* - co-first authors

Corresponding Author:
Megan L. Porter

Email: mlporter@hawaii.edu

Classification: BIOLOGICAL SCIENCES: Evolution

Keywords: Stomatopoda, opsin, retinal expression, *in situ* hybridization, evolution

Author Contributions

MLP & TWC conceptualized the project; MLP, HA, and MJB conducted all wet lab procedures; MLP conducted bioinformatics related to RNAseq data collection and analyses; MLP and HA performed all phylogenetic and evolutionary analyses. MLP and HA wrote the manuscript with input from all co-authors.

Competing interest statement: The authors declare no competing interests

This PDF file includes:

Main Text
Figures 1 to 4
Tables 1 to 2

1 **Abstract**

2 Stomatopod crustaceans possess some of the most complex animal visual systems, including at least 16
3 spectrally-distinct types of photoreceptive units (e.g. assemblages of photoreceptor cells). Here for the
4 first time we fully characterize the set of opsin genes expressed in retinal tissues and determine
5 expression patterns of each in the stomatopod *Neogonodactylus oerstedii*. Using a combination of
6 transcriptome and RACE sequencing, we identified 33 opsin transcripts expressed in each *N. oerstedii*
7 eye, which are predicted to form 20 long wavelength sensitive (LWS), 10 middle wavelength sensitive
8 (MWS), and three ultraviolet sensitive (UVS) visual pigments. Observed expression patterns of these
9 33 transcripts were highly unusual in five respects: (1) all long-wavelength and short/middle-
10 wavelength photoreceptive units expressed multiple opsins while ultraviolet photoreceptor cells
11 expressed single opsins; (2) most of the long-wavelength photoreceptive units expressed at least one
12 MWS opsin transcript; (3) the photoreceptors involved in spatial, motion, and polarization vision
13 expressed more transcripts than those involved in color vision; (4) there is a unique opsin transcript that
14 is expressed in all eight of the photoreceptive units devoted to color vision; and (5) expression patterns
15 in the peripheral hemispheres of the eyes suggest visual specializations not previously recognized in
16 stomatopods. Elucidating the expression patterns of all opsin transcripts expressed in the *N. oerstedii*
17 retina reveals the potential for previously undocumented functional diversity in the already complex
18 stomatopod eye and is a first step towards understanding the functional significance of the unusual
19 abundance of opsins found in many arthropod species' visual systems.

21 **Significance Statement**

22 Our study reveals the most elaborate opsin expression patterns ever described in any animal
23 eye. In mantis shrimp, a pugnacious crustacean renowned for its visual sophistication, we found
24 unexpected retinal expression patterns highlighting the potential for cryptic photoreceptor functional
25 diversity, including single photoreceptors that co-express opsins from different spectral clades and a

26 single opsin with a putative non-visual function important in color vision. This study demonstrates the
27 evolutionary potential for increasing visual system functional diversity through opsin gene duplication
28 and diversification, as well as changes in patterns of gene co-expression among photoreceptors and
29 retinula cells. These results have significant implications for the function of other visual systems,
30 particularly in arthropods where large numbers of retinally expressed opsins have been documented.

31 **Introduction**

32 Stomatopod crustaceans are successful predators in marine habitats. They are famous for their
33 aggressive nature, powerful prey capture and defense system, and in particular their highly unusual
34 visual system (1). Unlike other arthropods, each stomatopod compound eye consists of three distinct
35 regions – dorsal and ventral hemispheres separated by an equatorial region of specialized ommatidia
36 referred to as the ‘midband’. While the hemispheres are typically thought to have anatomically and
37 physiologically similar ommatidia throughout, the midband (MB) is most commonly formed from six
38 parallel rows of ommatidia that contain photoreceptors specialized for color (MB rows 1-4) and
39 polarization (MB rows 5 and 6) discrimination (**Fig. 1**) (2-5).

40 The Caribbean species *Neogonodactylus oerstedii* has been particularly well studied and is
41 thought to be representative of many stomatopods with six-row midbands (6-10). Altogether, its retina
42 is composed of 16 spectrally-distinct photoreceptive units (10). Similar to other crustaceans, each
43 ommatidium is composed of eight photoreceptor (retinula) cells (R1-8). In the hemispheres and in MB
44 rows 5 and 6, these eight cells are organized into two spectrally distinct layers composed of seven
45 retinula cells (R1-7) situated proximal to a single-celled eighth retinula (R8) photoreceptor. Ommatidia
46 in both hemispheres of *N. oerstedii* and in the two most ventral ommatidial rows of the midband, MB
47 rows 5 and 6, have two photoreceptive layers – an ultraviolet R8 receptor distal to a middle
48 wavelength-sensitive layer composed of R1-7 cells (6, 7, 9-11),

49 Most of the spectral diversity found in stomatopod retinas exists in the four dorsal-most
50 ommatidial midband rows (MB rows 1-4). In MB rows 1-4, the seven proximal photoreceptor cells
51 (R1-7) are subdivided into two separate photoreceptive units, formed by sets of three or four retinula
52 cells. Therefore, in MB rows 1-4 each ommatidium is formed from three successive photoreceptive
53 layers comprising the distal ultraviolet R8 photoreceptor, and middle and proximal photoreceptive
54 layers sensitive to the visible spectrum (**Fig. 1C**). Each of the three photoreceptor layers in each of the
55 first four midband rows is spectrally distinct, adding twelve classes to the overall spectral diversity of

the *N. oerstedii* eye. The spectral diversity of the photoreceptors in these layers is the result of distinct visual pigment expression, filtering through the visual pigments of photoreceptors of successive layers, and additional optical filtering by photostable pigments in the crystalline cones and rhabdoms that further tune their spectral sensitivities (3-7, 10-13).

In all animal visual systems, photoreceptor spectral absorbance is due to the expression of light-absorbing visual pigments. These visual pigments are composed of an opsin protein covalently bound to a vitamin-A derived chromophore. The spectral diversity observed among different photoreceptors within an organism's visual system is generally based on the expression of different opsin proteins bound to the same chromophore. Crustacean opsin genes have been divided into three distinct phylogenetic groups based on general absorbance properties of the visual pigments: long wavelength-sensitive (LWS), middle wavelength-sensitive (MWS), and short wavelength- or ultraviolet-sensitive (SWS/UVS) (14). Given the sixteen spectrally-distinct classes of photoreceptive units in *N. oerstedii* retina, a reasonable hypothesis is that there is one visual opsin expressed for each, for a total of sixteen opsins. However, previous studies found 15 different LWS opsin transcripts expressed in *N. oerstedii* retinas, which contain only six LWS photoreceptive units, suggesting the potential for rampant co-expression of opsin genes in single photoreceptors (8). Recent studies of *N. oerstedii* UVS opsins found many fewer transcripts expressed than spectrally-distinct R8 photoreceptor types (7). These differences between the number of opsins expressed and number of spectral types of photoreceptors (many more LWS opsins than long-wavelength photoreceptive units and fewer UVS opsins than ultraviolet photoreceptors), implies that photoreceptors and opsin proteins tuned to different parts of the visible spectrum may be under different evolutionary pressures.

The aim of this study is to characterize the set of opsin transcripts expressed in *Neogonodactylus oerstedii* retinal photoreceptors. Furthermore, due to the discrepancy between opsin diversity and photoreceptor diversity within a specific wavelength class, we examine the expression patterns of all identified opsin transcripts. This study provides the first fully characterized set of opsins

from a stomatopod compound eye and demonstrates that stomatopod visual systems exhibit extreme complexity both at the level of photoreceptor diversity and at the level of opsin expression patterns.

Results

Neogonodactylus oerstedii retinal opsin transcript diversity

In addition to previous results, we identified fifteen new *N. oerstedii* opsins, bringing the expressed retinal visual opsin total to 33 transcripts (7, 8). Of these 33 visual opsins, 24 are full-length transcripts, 8 are partial transcripts missing only the 5' end of the coding sequence, and one is a partial transcript missing coding regions on both the 5' and 3' ends. All 33 transcripts contain typical opsin structural motifs including the predicted seven-transmembrane domains, a lysine residue that forms a protonated Schiff base with the bound retinal chromophore, a glutamic acid residue at the predicted R-type opsin counterion site for the positive charge of the protonated lysine, and a conserved amino acid motif (e.g. glutamic acid / aspartic acid – arginine – tyrosine) in transmembrane-helix 3 that stabilizes the inactive state conformation of class A rhodopsin-like GPCRs (15).

In phylogenetic reconstructions including previously published crustacean opsin sequences, all *N. oerstedii* retinal opsins clustered into the expected three arthropod visual pigment clades (LWS, MWS and SWS/UVS; **Fig. 2**). In total, the 33 *N. oerstedii* retinal opsins consist of 20 LWS (5 newly identified), 10 MWS (all newly identified), and 3 UVS transcripts (previously identified (7)). Within each spectral clade the opsin transcripts were further divided into well-supported subgroups. The LWS opsins all clustered into the six previously identified stomatopod opsin subgroups A-F (8). The five new LWS opsins identified in this study belong to subgroup B (NoL18), subgroup C (NoL20), and subgroup E (NoL16, NoL17, and NoL19). In particular, the three opsin transcripts in LWS subgroup E represent the first characterized *N. oerstedii* sequences from this subgroup. For the MWS clade, the 10 opsins form three subgroups labeled G through I (**Fig. 2**). We assigned previously described *N. oerstedii* UVS opsins to new subgroups J (NoUV1 and NoUV3) and K (NoUV2).

106

107 *Neogonodactylus oerstedii opsin transcript retinal expression patterns*

108 All identified opsins appear to be functional based on conserved structural motifs, so we determined the
109 retinal expression pattern of each transcript (**Fig. 3**). We will describe the expression patterns of each
110 transcript first by comparing patterns within opsin spectral classes and then by patterns observed based
111 on photoreceptors with similar function or visual task (e.g. spatial and motion vision, color vision,
112 polarization vision, or UV vision).

113

114 *LWS opsin expression patterns:* Nineteen of the 20 LWS opsin transcripts were expressed
115 exclusively in long-wavelength photoreceptors (i.e. the R1-7 photoreceptors of the middle and
116 proximal tiers in MB rows 2, 3, 5 and 6, and both hemispheres). Subgroup A and B opsin transcripts
117 were exclusively expressed in photoreceptive layers composed of R1-7 cells of the hemispheres and
118 midband rows 5 and 6, although there was not a strict division of expression patterns by subgroup
119 (**Fig. 3A**). The three most similar transcripts in LWS subgroup A (NoL6, 7, 12) were all expressed in
120 midband rows 5 and 6. The fourth transcript in subgroup A (NoL8), however, was expressed in the R1-
121 7 photoreceptive units of both the midband rows 5 and 6 and in a subset of the hemisphere ommatidia,
122 including only those closest to the midband (**Fig. 3A**). In subgroup B, the single probe designed for
123 five nearly identical transcripts (NoL3-5, NoL10-11) showed that as a set, these were expressed in all
124 of the R1-7 photoreceptive layers of the hemisphere ommatidia; however, we cannot determine if there
125 are expression level differences between individual transcripts. The sixth transcript of this subgroup
126 (NoL18) was expressed only in the R1-7 photoreceptive layer of midband rows 5 and 6 (**Fig. 3A**).

127 The remaining transcripts in LWS subgroups D through F were mainly expressed in the long-
128 wavelength middle and proximal layers of midband rows 2 and 3 (**Fig. 3A**). However, the two
129 subgroup C opsin transcripts (NoL9, NoL20) showed more complex retinal expression patterns. NoL9
130 had the most unusual expression patterns of all of the LWS opsins, being consistently expressed in the

131 middle and proximal photoreceptive units of all of the midband rows used in color vision (e.g. rows 1-
132 4), including photoreceptors in rows 1 and 4 that have short-wavelength sensitivities (**Fig. 3A**). NoL20
133 had expression patterns that were difficult to interpret, with strong labeling in the row 3 proximal tier
134 but also weaker labeling in the row 3 middle and row 2 proximal tiers (**Fig. 3A**). In each of subgroups
135 D, E, and F, the transcripts had comparable expression patterns. In subgroup D, the highly similar
136 transcripts NoL1 and NoL2 were both expressed in the MB row 2 proximal tier (**Fig. 3A**), while two
137 subgroup E transcripts (NoL16 and NoL17) were expressed in the MB row 2 middle tier (**Fig. 3A**). The
138 third transcript in subgroup E, NoL19, did not clearly label any retinal photoreceptors, although it had
139 relatively high expression levels based on eye transcriptome estimates (2.4×10^4 TPM). Finally, two
140 subgroup F transcripts were expressed in midband row 3 photoreceptive layers, with NoL15 expressed
141 in the proximal tier and NoL13 in the middle tier. Uniquely, the third transcript in subgroup F, NoL14,
142 had a similar expression pattern to subgroup B transcripts and was expressed in all of the R1-7
143 photoreceptive units in the hemispheres of the eye.

145 *MWS opsin expression patterns:* The ten *N. oerstedii* MWS opsin transcripts exhibited two
146 general classes of expression pattern (**Fig. 3B**). The first was, as expected, expression in single
147 short/middle-wavelength photoreceptive units in the midband (i.e. the middle and proximal tiers of
148 rows 1 and 4). The second, unexpected, pattern of expression was for MWS transcripts found in
149 distinct regionalized sets of photoreceptors in the hemispheres that were previously assumed to be
150 sensitive to long-wavelengths of light. Each of the four transcripts in MWS subgroup G had a different
151 expression pattern. Three of the transcripts were expressed in short/middle-wavelength photoreceptive
152 layers: NoM1 in the midband row 1 middle tier, NoM3 in the row 1 proximal tier, and NoM4 in the
153 row 4 middle tier. The final subgroup G transcript, NoM2, was expressed only in the R1-7
154 photoreceptive unit of ommatidia that line the dorsal border of the eye (mDH, **Fig. 3B**).

155 The various transcripts in subgroup H had very different expression patterns, but all were

expressed in photoreceptors previously characterized as LWS. Transcript NoM10 was expressed in the R1-7 layer of midband rows 5 and 6, NoM8 in both dorsal and ventral hemisphere R1-7 photoreceptive units (similar to LWS transcript NoL8), and NoM7 and NoM9 in only the R1-7 layer of the ommatidia that line the dorsal border of the eye (**Fig. 3B**). Finally, in subgroup I, two similar transcripts shared expression in midband row 4 proximal layers containing photoreceptors with short/middle-wavelength sensitivity (**Fig. 3B**).

SWS opsin expression patterns: The expression patterns of the three transcripts found in the SWS opsin clade were published previously (7). Two of the three SWS transcripts were expressed in the ultraviolet R8 cells of the retina, with a nearly non-overlapping pattern. Transcript NoUV1 (subgroup J) was expressed in R8 cells of the dorsal hemisphere, midband rows 2-6, and the ventral hemisphere ommatidia immediately adjacent to the midband. NoUV2 (subgroup K) was found in the R8 cells of midband row 1 and the ventral hemisphere (**Fig. 3C**). NoUV3 (subgroup J) was not expressed at significant levels in any retinal photoreceptors, and also had no detectable expression in the transcriptome dataset.

Overall opsin expression in photoreceptor classes: All the MWS and LWS opsins apart from NoL19 (which based on *in situ* data had no detectable expression in the retina) were expressed in R1-7 cells (**Fig. 3**). Additionally, all of the LWS and MWS photoreceptive layers throughout the eye (both in hemispheres and in MB rows) expressed from two to nine opsins (**Fig. 4A; SI Appendix Fig. S2**). Comparisons of photoreceptive layer spectral sensitivities, spectral absorbances, and opsin expression patterns showed that those with similar spectral properties tend to express similar sets of opsins (**Fig. 4B**).

Differences were also found in expression patterns among photoreceptive units used for different visual tasks. In general, clusters of photoreceptors used for spatial, motion, and polarization

vision expressed opsins from multiple subgroups both within and across spectral clades (**Fig. 4A, B**). For example, the MB row 5 and 6 R1-7 layer, specialized for polarization vision, expressed a total of six transcripts from LWS subgroups A (4 transcripts) and B (1 transcript), and MWS subgroup H (1 transcript). The R1-7 layers in the hemispheres, used for spatial, motion, and polarization vision, express either six transcripts (6 LWS) in the most marginal ommatidia of the ventral hemisphere, eight transcripts (7 LWS, 1 MWS) in the dorsal and ventral hemisphere ommatidia other than those at the margins, or nine transcripts (6 LWS and 3 MWS) in the most marginal ommatidia of the dorsal hemisphere (**Figs. 3, 4**). In contrast, the photoreceptive layers in midband ommatidia devoted to color (e.g. row 1-4 middle and proximal tiers) generally expressed fewer opsins consisting of sets of highly similar transcripts from within the same spectral subgroup (e.g. LWS subgroup D transcripts NoL1 and NoL2 were both expressed in the midband row 2 proximal tier, LWS subgroup E transcripts were expressed in the MB row 2 middle tier, and MWS subgroup I opsins in the MB row 4 proximal tier). One notable exception to this pattern is transcript NoL9, which was expressed in all eight of the middle and proximal tiers in MB rows 1-4 (**Fig. 4A**). Finally, in contrast to long-wavelength and short/middle-wavelength photoreceptive units that express multiple opsins, ultraviolet R8 photoreceptors all expressed a single opsin transcript (**Fig. 3C**) (7).

Evidence for site-specific selection

Amino acid sites near the visual pigment chromophore tend to affect spectral tuning. Through comparison of our *N. oerstedii* alignment with a recently published jumping spider rhodopsin protein structure (16), we identified 26 sites within 5 Å of the chromophore (**dataset S2**). Of these, six sites were invariable, including the chromophore binding site, the counterion site, and one of the cysteine residues involved in disulfide bridge formation. Substitutions in opsins of S-A or F-Y near the chromophore are known to produce spectral shifts in the peak absorbance of visual pigments in many species (17-19). In the identified 26 sites in *N. oerstedii* opsins, seven sites contained S-A or F-Y

substitutions that varied within or between spectral clades (**Table 1**). Using tests of selection based on dN/dS ratios (PAML, MEME), we identified a further ten sites under positive selection across the *N. oerstedii* opsin phylogeny; one site (NoL1#69) was identified by both tests (**SI Appendix, Table S4**). None of these sites was within 5 Å of the chromophore. Using PAML random sites models, there was evidence for positive selection across the *N. oerstedii* opsin phylogeny (M7 versus M8, $p < 0.004$; $\omega = 1.15$), with two sites identified as under selection (NoL#69 and NoL#166, $p \leq 0.05$). In addition to the potential spectral tuning sites close to the chromophore with S-A or F-Y substitutions, an additional 12 sites from the set of 36 identified from all three methods had variation that corresponded with opsin clade structure (**Table 2**). The remaining eleven sites from the 36 identified from all three methods had variation both within and across clades, or variation that was not significantly different in amino acid properties (e.g. I-L substitutions).

217

218 **DISCUSSION**

219 *N. oerstedii* retinal opsin transcript numbers

With three UVS, ten MWS, and 20 LWS transcripts, *N. oerstedii* expresses an astounding diversity of visual opsins. The high numbers of characterized opsin transcripts are similar to numbers published in transcriptome estimates for other stomatopod species, specifically *Hemisquilla californiensis* (11 LWS, 2 MWS, and 1 SWS) and *Pseudosquilla ciliata* (15 LWS, 16 MWS, and 2 SWS) (14). Dragonflies and damselflies (Odonata) also have similar patterns of diversity, generally including 1 UVS, 1-8 insect SWS, and 8-21 LWS opsin genes (20), with different species expressing anywhere from 4 to 30 visual opsin genes (20-22). Genomic sequencing in many arthropod taxa has revealed unexpectedly large numbers of potential visual opsins, including 12 opsins in the mosquito *Anopheles gambia*, 9 of which are in visual opsin clades (23); 18 opsins in the horseshoe crab *Limulus polyphemus*, 13 of which are expressed in eyes (24); 46 opsins in the crustacean *Daphnia pulex*, 27 of which are likely used for vision (25); and 42 visual opsins (33 LWS, 7 MWS, and 2 SWS) in the penaid shrimp *Litopenaeus*

231 *vannamei* (26). Understanding the functional implications of these large numbers of opsin genes in
232 arthropods requires studies that link genomic opsin diversity to retinal expression patterns, as well as
233 research on actual photoreceptor spectral and polarization responses.

234 In arthropods where an abundance of expressed opsins has been found (i.e. horseshoe crabs,
235 odonates, stomatopods), most gene duplications occur in the LWS opsin clade, with fewer in the MWS
236 clade and very few in the SWS/UVS clade. These differences in frequency of gene duplication suggest
237 that different selection pressures act on visual pigments sensitive to different parts of the spectrum.
238 The stomatopod and odonate visual systems are among the most extreme examples of LWS opsin gene
239 duplication yet characterized, with *N. oerstedii* expressing 20 LWS and the dragonfly *Anax parthenope*
240 expressing 21 LWS opsins (20). What drives LWS opsin gene expansion in these taxa? In odonates,
241 Suvorov *et al.* (22) found that a heterozygote advantage model may explain the opsin gene repertoire
242 expansion, where positive selection acts on alleles that are subsequently duplicated and fixed in the
243 genome. Other possible explanations include genome characteristics and similarities in ecologies
244 (stomatopods and odonates are both highly visual predators).

245 The high numbers of MWS transcripts found in *N. oerstedii* (10 opsins) are in line with the
246 numbers of MWS genes revealed in other crustacean studies (*Daphnia pulex* genome 25 MWS genes
247 (25), *Litopenaeus vannamei* genome 7 MWS genes (26)). Unlike these crustacean examples, one of the
248 few published chelicerate genomes, from *L. polyphemus*, contains a single MWS opsin (24).
249 Interestingly, among stomatopod species the expressed number of MWS opsin transcripts appears to
250 vary strongly, with only 3 transcripts recovered in *H. californiensis*, compared to 10 in *N. oerstedii* and
251 16 in *P. ciliata* (this study; 14).

252 In insects, blue-sensitive opsins are an evolutionarily distinct lineage from the crustacean and
253 horseshoe crab MWS opsins, clustering with what we call here the SWS clade; therefore, the 10 *N.*
254 *oerstedii* MWS opsins and the 8 MWS opsins in *A. parthenope* that form putatively short/middle-
255 wavelength sensitive visual pigments must be an independent set of duplications. Although most

orders of insects do not appear to contain sequences closely related to the crustacean MWS clade, recent studies found crustacean-type MWS opsins in early branching hexapod lineages (*Sminthurus viridis*, Collembola; *Machilis hrabei*, Archaeognatha) and myriapods (*Scutigera coleoptrata*) (27).

N. oerstedii retinal opsin expression patterns

Only a few genomic and transcriptomic studies of arthropod visual systems have determined the expression pattern of each identified transcript. In this study we demonstrated retinal expression patterns for 31 of the 33 described opsin transcripts. The two opsin transcripts, NoL19 and NoUV3, with no clear retinal expression pattern may be expressed in parts of the nervous system that we did not include in our study. This idea is supported for NoL19 by two pieces of information. First, NoL19 had expression levels (2.4×10^4 TPM) in the transcriptome similar to other LWS opsin transcripts. Second, a recent study by Donohue *et al.* (28) found NoL19 (identical to transcript NoT27696g1 in (28)) to be expressed in the cerebral ganglion. The case for NoUV3 is less clear, as it had no detectable expression in the transcriptome dataset; this may be due to expression in a very small subset of cells or the possibility that the tissue of expression was not included in the generation of the transcriptome data. Future studies should check the visual neuropils for both of these “orphan” eye transcripts and the CNS for expression of NoUV3.

Determining the retinal expression patterns for all 33 opsins found in the *N. oerstedii* retina is a major contribution to understanding the function of the complex stomatopod visual system and, more broadly, other arthropod eyes that express a superabundance of opsin transcripts relative to described photoreceptor types. Perhaps the most comprehensive previous study of arthropod opsin expression patterns is that of Battelle *et al.* (24), which documented 18 opsin genes in the *L. polyphemus* genome and determined the expression patterns of each across the visual system and CNS. Thirteen of these opsins are expressed in ocular tissues of the median, ventral or lateral eyes, although only 10 of them are found in what are typically considered visual opsin clades. *L. polyphemus* eyes all express opsins

281 from multiple spectral clades: the median eyes express 7 LWS, 1 MWS, and 1 SWS opsins, while the
282 ventral and lateral eyes each express 4 LWS, 1 MWS, and 1 SWS opsins (24). A similar pattern was
283 found in *N. oerstedii* eyes, where all of the R1-7 photoreceptive layers in MB rows 5 and 6 and the
284 hemispheres (with the exception of the most marginal region of the ventral hemisphere) express
285 numerous LWS opsin transcripts as well as at least one MWS opsin transcript.

286 Given that the long-wavelength and short/middle-wavelength photoreceptive units in the *N.*
287 *oerstedii* retina express multiple opsin transcripts, an important question for study is how these opsins
288 are distributed among the retinula cells that contribute to each unit. In the ommatidia of the *N. oerstedii*
289 hemispheres, the R1-7 photoreceptive layer is composed of seven cells which form a fused central
290 rhabdom. Determining which sets of the R1-7 cells express MWS opsin transcripts in the LWS
291 photoreceptive units may provide further clues to the functional significance of these apparently
292 spectrally mismatched visual pigments. Based on LWS and MWS opsin expression patterns across
293 hemisphere ommatidia, there are more opsins expressed (6 LWS and 3 MWS opsins in the marginal
294 dorsal hemisphere ommatidia; 7 LWS and 1 MWS in the dorsal and ventral hemisphere ommatidia)
295 than the number of previously characterized spectrally distinct photoreceptive units (1 LWS, 0 MWS).
296 Because each of the hemisphere ommatidia contains up to seven photoreceptor cells (R1-7)
297 contributing to LWS sensitivity, this pattern of 8 or 9 expressed LWS/MWS opsins per ommatidia
298 suggests that some photoreceptor cells in hemisphere ommatidia must express multiple opsin
299 transcripts. However, we cannot predict at this point whether all retinula cells express all opsin
300 transcripts, to what degree subsets of these cells express subsets of opsins, or whether all of the
301 transcripts identified are translated into stable proteins capable of binding chromophore. Co-expression
302 of multiple opsins in arthropod photoreceptor cells has been shown in a number of species, including
303 butterflies (29, 30), mosquitoes (31, 32), horseshoe crabs (24), and fiddler crabs (33). In insects (e.g.
304 butterflies, mosquitoes), opsin co-expression has been shown to broaden the spectral sensitivity of
305 photoreceptors (29) or enhance vision in low light environments (31). In studies of opsin co-expression

306 in crustaceans, Rajkumar *et al.* (33) found that the two MWS opsins expressed in R1-7 layer of the
307 fiddler crab *Uca pugilator* were co-expressed in only one of the seven retinula cells forming the
308 rhabdom, while each of the remaining six retinula cells expressed only one of the two possible
309 transcripts. Based on this expression pattern involving only two opsin transcripts, the expression
310 patterns at the level of individual retinula cells in stomatopods is likely to be complex, and possibly
311 unique for each opsin. In one of the few published studies on stomatopod opsin expression patterns, at
312 least two LWS opsin transcripts were co-expressed in all retinula cells of the retina of *Squilla empusa*
313 (34), a species with a relatively simple eye having only two rows of ommatidia in the midband. In the
314 more complicated 6-midband row eye of *N. oerstedii*, if each retinula cell (R1-7) expressed a unique
315 set of opsins, then photoreceptive units composed of fewer cells (three or four cells in midband
316 photoreceptor tiers versus seven in hemisphere R1-7 photoreceptive units) would be predicted to
317 express fewer transcripts. Indeed, we observe this pattern in the *N. oerstedii* retina, where tiered middle
318 and proximal photoreceptive layers in the midband express fewer opsin transcripts than the hemisphere
319 R1-7 layer (**Fig. 4**).

320 Our data also illustrate how molecular studies can be used to reveal new photoreceptive
321 structures missed by other methods. An example in *N. oerstedii* is the very marginal region of the
322 dorsal hemisphere, containing ommatidia that line the dorsal border of the eye. The expression of three
323 different MWS opsin transcripts in this region is reminiscent of the ‘dorsal rim’ area of insect eyes
324 typically used to detect shorter-wavelength skylight polarization patterns (35). Histological sections
325 used for *in situ* hybridization, although not ideal for anatomical studies, suggest comparable
326 morphological changes in the dorsal-most part of the *N. oerstedii* eye, where all the photoreceptors are
327 directed upward. This specialized region is the first identified in a crustacean eye. In insects,
328 photoreceptors in the dorsal rim area have been found to be sensitive to either UV, blue, or green
329 wavelengths of light and may exhibit a suite of morphological changes, including altered or lost lens
330 structures, orthogonally arranged microvilli, loss of screening pigment, an upward directed optical axis,

and significantly reduced rhabdom lengths (35). Consistent with what is known from insect eyes, our molecular data suggest the presence of blue-green to blue visual pigments in the dorsal-most eye region of stomatopods. Previous studies of related species suggest this region of the eye is also unique with respect to the presence of UV filters (11). Investigations of ommatidia in the stomatopod eye dorsal rim region are needed, to understand whether or not photoreceptor anatomy and physiology of this region differs from that of the adjacent hemisphere receptors and to learn if *N. oerstedii* show behavioral evidence of using skylight polarization patterns for navigation underwater.

Why do stomatopods have so many opsins in their retinas?

The retinal expression patterns of identified opsin transcripts were far more complex than we predicted. Photoreceptive layers formed from retinula cells R1-7 in *N. oerstedii* express as many as nine different opsin transcripts (**Fig. 4**). This is particularly surprising because absorbance spectra previously measured by microspectrophotometry (MSP) suggest that each photoreceptive unit, with the possible exception of those of the hemispheres, possesses only a single spectral class of visual pigment (10). However, absorbance spectral data are unsuitable for detecting the presence of multiple visual pigments if they have similar absorption spectra or if one or more components are present in relatively low concentration.

When multiple opsins are expressed in single photoreceptors, the mixture's presence is often obvious from the broadening of the spectral absorption or sensitivity of the receptor. In vertebrates, for example cichlid fish or nocturnal rodents, the absorption effects of visual pigment coexpression have generally been interpreted as adaptations to the spectral or luminosity features of the environment (36-38). In arthropods the situation is often more puzzling. The best data available are from butterflies, where opsin coexpression is common in subsets of retinal photoreceptors that have also been physiologically characterized. In some cases, the mixture produces spectrally obvious effects, forming receptors with extremely broad, or even double, spectral peaks (29, 30, 39). In others, however,

356 photoreceptors with mixed opsins have spectral sensitivities similar to receptors with only one
357 rhodopsin class (30, 40). Similar to our MSP findings in *N. oerstedii*, previous crustacean studies on
358 opsin coexpression in an intertidal grapsid crab (41) and a deep-sea lophogastrid (42) found that
359 spectral sensitivity of photoreceptors was explainable by a single visual pigment in the retina despite
360 the mix of expressed opsins. Electrophysiological recordings from single photoreceptor cells in the
361 color receptors (i.e. from photoreceptive units in the dorsal four midband rows) of *N. oerstedii* indicate
362 narrow-band sensitivity (43), essentially identical to predictions made based on models using single
363 visual pigments plus the overlying filters (10). This is not particularly surprising, given that these
364 receptors express few opsins and are heavily filtered. On the other hand, physiological measurements
365 from receptors in the hemispheres, which are not filtered and which express up to nine opsin
366 transcripts, also find spectral sensitivities explainable by the presence of a single visual pigment (44).
367 Given the consistency of these results from multiple crustaceans, it appears that the multiple opsin
368 expression we report here rarely (or never) leads to spectral sensitivity broadening in these animals.

369 One possible explanation of the seeming contradiction between the complex opsin expression
370 results and the MSP measurements is that some transcribed opsins, for example the LWS-subgroup C
371 opsins expressed in all of the visible light photoreceptors in MB rows 1-4, are either untranslated,
372 nonfunctional as proteins, or have a yet unidentified non-visual function ('functionality' hypothesis).
373 Alternatively, only one (or at least a very few) of the expressed opsins dominates in visual pigments in
374 the rhabdom of a given photoreceptive layer ('dominance' hypothesis). A third possibility is that opsins
375 belonging to the same subgroup (i.e. A, B, C, etc.) that are expressed in the same cluster of
376 photoreceptors may have similar or identical absorption spectra ('similarity' hypothesis). A final
377 possibility is that some of the expressed visual opsins have non-visual functions within the retina ('non-
378 visual' hypothesis).

379 Although these hypotheses are not mutually exclusive, and the expression patterns of different
380 transcripts may be explained by alternate hypotheses, our data allow us to investigate these three

alternatives using several approaches: 1) analyzing transcript sequences for sites important to visual pigment function (functionality hypothesis), 2) investigating expression levels of selected transcripts expressed in similar sets of photoreceptors (dominance hypothesis), and 3) inspecting the residue identities for sites under positive selection and potential spectral tuning sites within groups of related sequences expressed in the same sets of photoreceptors (similarity hypothesis). For hypotheses of some transcripts having non-visual functions, we can only speculate as to the potential roles based on previous studies in other arthropods.

388

Functionality Hypothesis: All 33 transcripts characterized here contained typical opsin structural motifs, including the chromophore binding and counterion sites, and can therefore be translated into functional proteins. Within the *N. oerstedii* opsins a set of 10 sets of sites were identified that were under positive selection (**SI Appendix, Table S4**). Because most of these sites are far from the chromophore binding pocket, it is likely that variations at these sites affect overall opsin function and GPCR signaling, rather than spectral tuning (34). Although many of these sites have amino acid substitutions that vary across or within clades (**Table 2**), at this point it is difficult to determine how this amino acid variation may affect overall visual pigment function and thus the veracity of the functionality hypothesis. Future studies of protein localization will help identify transcripts translated into functional proteins, while the sites under positive selection may serve as candidates for *in vitro* studies of arthropod opsin functional diversity unrelated to spectral tuning.

400

Dominance Hypothesis: As a first approach to investigating whether single transcripts dominate in a particular photoreceptive unit, we considered both the relative level of labeling in *in situ* hybridization sections and the calculated retinal expression levels of co-expressed sets of opsin transcripts. Although this is a poor measure of abundance, some sets of transcripts expressed in the same photoreceptive layer all had strong *in situ* labeling, suggesting significant expression levels for multiple transcripts

(e.g. the expression patterns of transcripts in LWS subgroups A and B in MB rows 5 and 6, or NoL1 and NoL2 expression in the MB row 2 proximal tier; **Fig. 3A**). Due to complex and overlapping expression patterns in multiple types of photoreceptive units, we have chosen not to present estimated expression levels of each transcript here because they are difficult to interpret. However, there are several sets of LWS transcripts with similar expression patterns, making estimated expression levels comparable. For example, transcripts from two different subgroups (e.g. NoL6-7 and NoL12 from Group A, and NoL18 from Group B) that have the same expression pattern in only the R1-7 layer of MB rows 5 and 6 all had similar levels of expression (NoL6 = 1.2×10^5 TPM; NoL7 = 1.4×10^5 TPM; NoL12 = 1.5×10^5 TPM; NoL18 = 1.0×10^5 TPM), illustrating that at least for some photoreceptive units there is not a single dominant expressed opsin. Similarly, the closely related sets of transcripts expressed in the MB row 2 proximal tier (NoL1 = 2.8×10^4 TPM; NoL2 = 2.5×10^4 TPM) and the row 2 middle tier (NoL16 = 1.3×10^4 TPM; NoL17 = 4.0×10^3 TPM) all had similar levels of expression (see **Fig. 3A**). Together, although neither of these methods can specifically test this question, results suggest that in most stomatopod retinal photoreceptive layers there is probably not a single dominant visual pigment responsible for the previously measured spectral absorbances, and therefore we consider the dominance hypothesis an unlikely explanation for stomatopod retinas. However, to establish this conclusion definitively, methods for determining translation efficiency as well as visual pigment functionality are needed.

Similarity Hypothesis: The similarity hypothesis, at its core, addresses the fate of opsin gene duplicates in the stomatopod genome. There is a large literature on evolutionary models for the maintenance of gene duplicates (45), distilling down to three outcomes: 1) the evolution of a new function ('neofunctionalization' - 46); 2) the division of ancestral functions among duplicates ('subfunctionalization'); 3) the conservation of function in both duplicates ('gene conservation'). In the context of stomatopod opsin genes that encode visual pigments, gene duplication followed by

431 subfunctionalization and/or neofunctionalization has led to the large set of visual pigments with diverse
432 peak spectral absorbances expressed in distinct photoreceptor types. However, what processes are at
433 play for duplicated genes expressed in the same photoreceptor is still an open question. There is
434 evidence from *S. empusa* that co-expressed, closely-related gene variants may interact differently with
435 the phototransduction machinery (34) and thus may tune photoreceptor cell temporal responses.

436 The only potential evolutionary model that we can address here, given our data for opsin gene
437 duplicates expressed within a single photoreceptive structure, is gene conservation. Mechanisms for
438 maintaining a gene duplicate include dosage models, where producing more of a gene product provides
439 a functional advantage. To address whether related sets of opsin transcripts may form visual pigments
440 with similar peak absorbances, we compared the residues at predicted spectral tuning sites across
441 transcript sequences that were closely related and also expressed in the same set of photoreceptor cells.
442 For example, the MB row 2 proximal and middle tiers both express two opsin transcripts that are
443 evolutionarily closely related (**Fig. 3A**), and that do not vary in amino acid at any of the potential
444 spectral tuning sites (e.g. subgroup E NoL16 and NoL17 in the MB row 2 middle tier and subgroup D
445 NoL1 and NoL2 in the MB row 2 proximal tier; **Table 1**). This suggests that spectral absorbance of the
446 encoded visual pigments does not vary, at least for these sets of transcripts, and they therefore may
447 adhere to the dosage model of gene duplicate conservation. Another possible factor in gene
448 conservation is heterodimerization, particularly in the LWS polarization-sensitive photoreceptive layers
449 in the hemispheres and MB rows 5 and 6, which all express 6 or more opsin transcripts (**Figs. 4A, B**).
450 Large numbers of expressed transcripts with similar spectral properties may play a role in increased
451 dimerization and/or heterodimerization and improved chromophore orientation along the microvillar
452 axis, improving the polarization sensitivity of these photoreceptors (47-49).

453

454 *Non-visual Hypothesis:* Although we have no way to test this hypothesis given the data presented here,
455 previous studies of *Drosophila* have shown that visual opsins can have light-independent functions,

456 including thermal preference (50), mechanoreception (51),(52) and larval locomotion. Therefore, we
457 cannot rule out the possibility that some of the opsin transcripts expressed in the *N. oerstedii* retina
458 have non-visual functions. Further studies are needed to begin to sort out these potential functions.
459

460 **Summary**

461 Elucidating the expression patterns of all 33 opsin transcripts found in *N. oerstedii* retinas is a major
462 step towards understanding the functional significance of the large abundance of opsins described from
463 many arthropod species' visual systems. This study demonstrates that the already complex
464 photoreceptor diversity of the stomatopod visual system is even more complex at the molecular level of
465 phototransduction. The observed expression patterns, particularly for the MWS opsins, were
466 unexpected and revealed the potential for previously unsuspected functional diversity in the
467 photoreceptors of the already complex stomatopod eye. In the retina of *N. oerstedii*, we show that all
468 long wavelength- and short/middle wavelength-sensitive photoreceptive layers express multiple opsins,
469 that the R1-7 photoreceptive units of the hemispheres and midband rows 5 and 6 express more
470 transcripts than any of the middle or proximal tiers involved in color vision in midband rows 1-4, and
471 that all of the hemisphere and midband row 5 and 6 long wavelength-sensitive photoreceptive units
472 express at least one MWS opsin transcript. In addition, there is a single opsin transcript that is uniquely
473 expressed in all eight of the midband row 1-4 middle and proximal tiers devoted to color vision.
474

475 **Materials & Methods**

476 *Data Availability*

477 The generated *N. oerstedii* transcriptome and opsin transcript sequences are available at Genbank (SRA
478 PRJNA609025, Accession numbers MT112859-MT112888). The full details of our methods, including
479 sequence alignments, are presented in the SI Appendix and SI datasets.
480

481 *RNAseq and opsin transcript identification*

482 Total RNA was isolated using TRIzol (Life Technologies) and purified using a RNeasy kit (Qiagen). A
483 transcriptome generated and assembled by Cofactor Genomics was queried using local BLAST+ (ver.
484 2.2.28) to identify opsin contigs. Identified opsin contigs, as well as previously published *N. oerstedii*
485 opsin sequences (7, 8), were confirmed using transcript-specific primers in RT-PCR, while degenerate
486 primers were used to search for possible transcripts missed by all other methods. RACE methods were
487 used to sequence the ends of incomplete transcripts (8).

488

489 *Opsin phylogenetic analysis*

490 Each opsin transcript was translated and aligned with other opsin sequences using MUSCLE (53, 54) as
491 implemented by MEGA 6.06 (55) to create two datasets: 1) *N. oerstedii* transcripts with other
492 published arthropod opsin sequences; and 2) *N. oerstedii* opsins only to use for mapping ISH results.
493 Phylogenetic analyses were carried out using maximum likelihood methods executed in RAxML 8.1.5
494 (56-58), with the best-fit model as selected by ProtTest3 (59, 60). Additionally, posterior probabilities
495 using Bayesian inference methods coupled with a Markov chain Monte Carlo algorithm executed in
496 MrBayes 3.2.6 (61, 62) were generated on the CIPRES Science Gateway v3.3 (63, 64).

497

498 *Analyses of spectral tuning sites, gene conversion, and selection*

499 SWISS-PDBVIEWER (65-67) was used to fit a *N. oerstedii* opsin (NoL1) to the jumping spider
500 rhodopsin protein structure (PDB:619K) (16), and to identify all residues predicted to be within five Å
501 of the chromophore. These candidate sites were screened based on how residue identities mapped to
502 spectral clades, specific amino acid properties, and correspondence with sites previously identified to
503 cause shifts in spectral absorbance of the visual pigment. Tests for selection across branches and sites
504 were accomplished using the CodeML program in the PAML software package (68, 69) as
505 implemented in Easycodeml (70), and the HyPhy software package on the Datamonkey webserver

506 (71-74). For residue identification, amino acids were numbered based on their position in the *N.*
507 *oerstedii* reference sequence (NoL1#), the *Hasarius adansoni* crystal structure (JSR1#), or bovine
508 rhodopsin (BvRh#).

509

510 *In situ hybridization*

511 Dissected eyes were fixed in 4% paraformaldehyde in 10mM phosphate buffer (pH 7.4), cryoprotected
512 with sucrose, permeated with Tissue-Tek O.C.T Compound (Sakura Finetek), quick-frozen, and
513 sectioned at a thickness of 12 μ m in a cryostat cooled to -20°C. With the exception of five transcripts
514 (NoL3-5, 10, and 11) with similar sequences, unique riboprobes were designed for each transcript
515 based on the 3' UTR of each opsin mRNA. Riboprobes were generated as described in (75).
516 Hybridizations were performed on cryosections as described in (75, 76). Individual hybridization
517 temperatures were calculated by an equation for RNA duplex formation derived by Bodkin and
518 Knudson (77).

519

520 **Acknowledgements**

521 This research was funded by grants from the National Science Foundation (IOS 0721608) and the Air
522 Force Office of Scientific Research through Grant Numbers FA9550-06-1-0117, FA9550-09-1-0149,
523 and FA9550-12-0321. Thanks to Mireille Steck for help with *in situ* hybridization and expression data.

References

1. Schram FR, *et al.* (2013) Subclass Hoplocarida Calman, 1904: Order Stomatopoda Latreille, 1817. *Treatise on Zoology - Anatomy, Taxonomy, Biology; The Crustacea; Revised and updated, as well as extended from the Traité de Zoologie*, eds Vaupel Klein JC, Charmantier-Daures M, & Schram FR (Brill, Leiden-Boston), Vol Volume 4, pp 179-356.
2. Marshall NJ (1988) A unique colour and polarization vision system in mantis shrimps. *Nature* 333(6173):557-560.
3. Marshall NJ, Land MF, King CA, & Cronin TW (1991) The compound eyes of mantis shrimps (Crustacea, Hoplocarida, Stomatopoda). I. Compound eye structure: the detection of polarized light. *Philosophical Transactions of the Royal Society B: Biological Sciences* 334(1269):33-56.
4. Marshall NJ, Land MF, King CA, & Cronin TW (1991) The compound eyes of mantis shrimps (Crustacea, Hoplocarida, Stomatopoda). II. Colour pigments in the eyes of stomatopod crustaceans: polychromatic vision by serial and lateral filtering. *Philos Trans R Soc Lond B Biol Sci* 334(1269):57-84.
5. Marshall J, Cronin TW, & Kleinlogel S (2007) Stomatopod eye structure and function: a review. *Arthropod Struct Dev* 36(4):420-448.
6. Marshall J & Oberwinkler J (1999) The colourful world of the mantis shrimp. *Nature* 401(6756):873-874.
7. Bok MJ, Porter ML, Place AR, & Cronin TW (2014) Biological sunscreens tune polychromatic ultraviolet vision in mantis shrimp. *Curr Biol* 24(14):1636-1642.
8. Porter ML, Bok MJ, Robinson PR, & Cronin TW (2009) Molecular diversity of visual pigments in Stomatopoda (Crustacea). *Visual neuroscience* 26(3):255-265.

- 548 9. Cronin TW, Marshall N, & Caldwell R (1996) Visual pigment diversity in two genera of mantis
549 shrimps implies rapid evolution (Crustacea; Stomatopoda). *Journal of Comparative Physiology*
550 *A: Sensory, Neural, and Behavioral Physiology* 179:371-384.
- 551 10. Cronin TW & Marshall NJ (1989) Multiple spectral classes of photoreceptors in the retinas of
552 gonodactyloid stomatopod crustaceans. *Journal of Comparative Physiology A* 166(2):261-275.
- 553 11. Bok MJ, Porter ML, & Cronin TW (2015) Ultraviolet filters in stomatopod crustaceans:
554 diversity, ecology and evolution. *J Exp Biol* 218(Pt 13):2055-2066.
- 555 12. Cronin TW, Bok MJ, Marshall NJ, & Caldwell RL (2014) Filtering and polychromatic vision in
556 mantis shrimps: themes in visible and ultraviolet vision. *Philos Trans R Soc Lond B Biol Sci*
557 369(1636):20130032.
- 558 13. Cronin TW & Marshall NJ (1989) A retina with at least ten spectral types of photoreceptors in a
559 mantis shrimp. *Nature* 339(6220):137-140.
- 560 14. Porter ML, *et al.* (2013) The evolution of complexity in the visual systems of stomatopods:
561 insights from transcriptomics. *Integr Comp Biol* 53(1):39-49.
- 562 15. Vogel R, *et al.* (2008) Functional role of the "ionic lock"--an interhelical hydrogen-bond
563 network in family A heptahelical receptors. *J Mol Biol* 380(4):648-655.
- 564 16. Varma N, *et al.* (2019) Crystal structure of jumping spider rhodopsin-1 as a light sensitive
565 GPCR. *Proc Natl Acad Sci U S A* 116(29):14547-14556.
- 566 17. Salcedo E, Farrell D, Zheng L, Phistry M, & Bagg E (2009) The Green-absorbing *Drosophila*
567 Rh6 Visual Pigment Contains a Blue-shifting *Journal of Biological*
- 568 18. Hope AJ, Partridge JC, Dulai KS, & Hunt DM (1997) Mechanisms of wavelength tuning in the
569 rod opsins of deep-sea fishes. *Proc Biol Sci* 264(1379):155-163.
- 570 19. Takahashi Y & Ebrey TG (2003) Molecular basis of spectral tuning in the newt short
571 wavelength sensitive visual pigment. *Biochemistry* 42(20):6025-6034.

- 572 20. Futahashi R, *et al.* (2015) Extraordinary diversity of visual opsin genes in dragonflies. *Proc*
573 *Natl Acad Sci U S A* 112(11):E1247-1256.
- 574 21. Chauhan P, *et al.* (2014) De novo transcriptome of *Ischnura elegans* provides insights into
575 sensory biology, colour and vision genes. *BMC Genomics* 15:808.
- 576 22. Suvorov A, *et al.* (2017) Opsins have evolved under the permanent heterozygote model:
577 insights from phylotranscriptomics of Odonata. *Mol Ecol* 26(5):1306-1322.
- 578 23. Hill CA, *et al.* (2002) G protein-coupled receptors in *Anopheles gambiae*. *Science*
579 298(5591):176-178.
- 580 24. Battelle BA, *et al.* (2016) Opsin Repertoire and Expression Patterns in Horseshoe Crabs:
581 Evidence from the Genome of *Limulus polyphemus* (Arthropoda: Chelicerata). *Genome Biol*
582 *Evol* 8(5):1571-1589.
- 583 25. Colbourne JK, *et al.* (2011) The ecoresponsive genome of *Daphnia pulex*. *Science*
584 331(6017):555-561.
- 585 26. Zhang X, *et al.* (2019) Penaeid shrimp genome provides insights into benthic adaptation and
586 frequent molting. *Nature communications* 10(1):356.
- 587 27. Henze MJ & Oakley TH (2015) The Dynamic Evolutionary History of Pancrustacean Eyes and
588 Opsins. *Integr Comp Biol* 55(5):830-842.
- 589 28. Donohue MW, Carleton KL, & Cronin TW (2017) Opsin Expression in the Central Nervous
590 System of the Mantis Shrimp *Neogonodactylus oerstedii*. *Biol Bull* 233(1):58-69.
- 591 29. Arikawa K, Mizuno S, Kinoshita M, & Stavenga DG (2003) Coexpression of two visual
592 pigments in a photoreceptor causes an abnormally broad spectral sensitivity in the eye of the
593 butterfly *Papilio xuthus*. *J Neurosci* 23(11):4527-4532.
- 594 30. Ogawa Y, *et al.* (2012) Coexpression of three middle wavelength-absorbing visual pigments in
595 sexually dimorphic photoreceptors of the butterfly *Colias erate*. *J Comp Physiol A Neuroethol*
596 *Sens Neural Behav Physiol* 198(12):857-867.

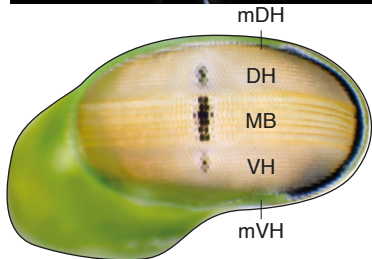
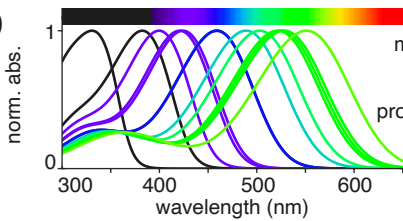
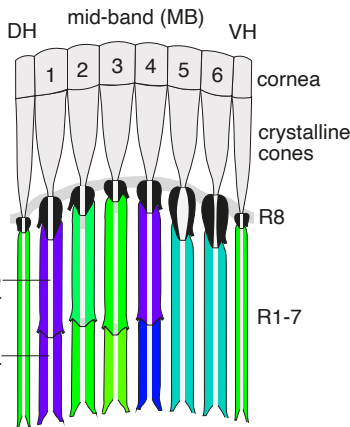
- 597 31. Hu X, Whaley MA, Stein MM, Mitchell BE, & O'Tousa JE (2011) Coexpression of spectrally
598 distinct rhodopsins in *Aedes aegypti* R7 photoreceptors. *PLoS One* 6(8):e23121.
- 599 32. Hu X, Leming MT, Whaley MA, & O'Tousa JE (2014) Rhodopsin coexpression in UV
600 photoreceptors of *Aedes aegypti* and *Anopheles gambiae* mosquitoes. *J Exp Biol* 217(Pt
601 6):1003-1008.
- 602 33. Rajkumar P, Rollmann SM, Cook TA, & Layne JE (2010) Molecular evidence for color
603 discrimination in the Atlantic sand fiddler crab, *Uca pugilator*. *Journal of Experimental Biology*
604 213(Pt 24):4240-4248.
- 605 34. Valdez-Lopez JC, *et al.* (2018) Sequence, Structure, and Expression of Opsins in the
606 Monochromatic Stomatopod *Squilla empusa*. *Integrative and Comparative Biology* 58(3):386-
607 397.
- 608 35. Labhart T & Meyer EP (1999) Detectors for polarized skylight in insects: a survey of
609 ommatidial specializations in the dorsal rim area of the compound eye. *Microsc Res Tech*
610 47(6):368-379.
- 611 36. Peichl L (2005) Diversity of mammalian photoreceptor properties: adaptations to habitat and
612 lifestyle? *Anat Rec A Discov Mol Cell Evol Biol* 287(1):1001-1012.
- 613 37. Dalton BE, Loew ER, Cronin TW, & Carleton KL (2014) Spectral tuning by opsin
614 coexpression in retinal regions that view different parts of the visual field. *Proc Biol Sci*
615 281(1797).
- 616 38. Dalton BE, Lu J, Leips J, Cronin TW, & Carleton KL (2015) Variable light environments
617 induce plastic spectral tuning by regional opsin coexpression in the African cichlid fish,
618 *Metriaclicma zebra*. *Mol Ecol* 24(16):4193-4204.
- 619 39. Chen P-J, Awata H, Matsushita A, Yang E-C, & Arikawa K (2016) Extreme spectral richness in
620 the eye of the common bluebottle *Graphium sarpedon*. *Frontiers in Ecology and Evolution*
621 4(18):1-18.

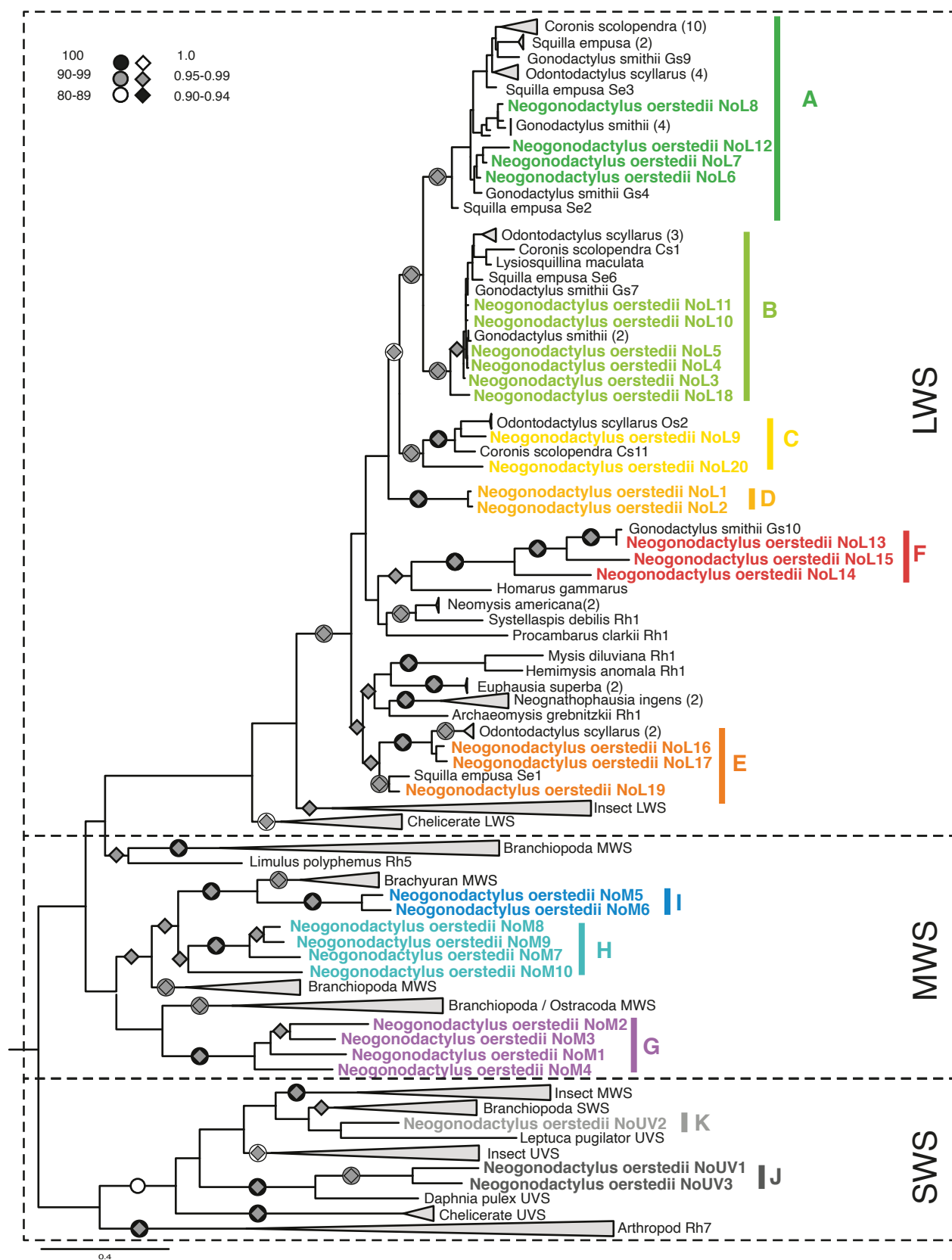
- 622 40. Kitamoto J, Sakamoto K, Ozaki K, Mishina Y, & Arikawa K (1998) Two visual pigments in a
623 single photoreceptor cell: identification and histological localization of three mRNAs encoding
624 visual pigment opsins in the retina of the butterfly *Papilio xuthus*. *J Exp Biol* 201(Pt 9):1255-
625 1261.
- 626 41. Sakamoto K, Hisatomi O, Tokunaga F, & Eguchi E (1996) Two opsins from the compound eye
627 of the crab *Hemigrapsus sanguineus*. *J Exp Biol* 199(Pt 2):441-450.
- 628 42. Frank TM, Porter M, & Cronin TW (2009) Spectral sensitivity, visual pigments and screening
629 pigments in two life history stages of the ontogenetic migrator *Gnathophausia ingens*. *Journal*
630 *of the Marine Biological Association of the United Kingdom* 89(01):119.
- 631 43. Cronin TW & Marshall J (2004) The unique visual world of mantis shrimps. *Complex worlds*
632 *from simpler nervous systems*. MIT Press, Cambridge, MA:239-268.
- 633 44. Cronin TW & King CA (1989) Spectral Sensitivity of Vision in the Mantis Shrimp,
634 *Gonodactylus oerstedii*, Determined Using Noninvasive Optical Techniques. *Biol Bull*
635 176(3):308-316.
- 636 45. Hahn MW (2009) Distinguishing among evolutionary models for the maintenance of gene
637 duplicates. *J Hered* 100(5):605-617.
- 638 46. Force A, *et al.* (1999) Preservation of duplicate genes by complementary, degenerative
639 mutations. *Genetics* 151(4):1531-1545.
- 640 47. Roberts NW, Porter ML, & Cronin TW (2011) The molecular basis of mechanisms underlying
641 polarization vision. *Philosophical Transactions of the Royal Society B: Biological Sciences*
642 366(1565):627-637.
- 643 48. Prinster SC, Hague C, & Hall RA (2005) Heterodimerization of G protein-coupled receptors:
644 specificity and functional significance. *Pharmacol Rev* 57(3):289-298.

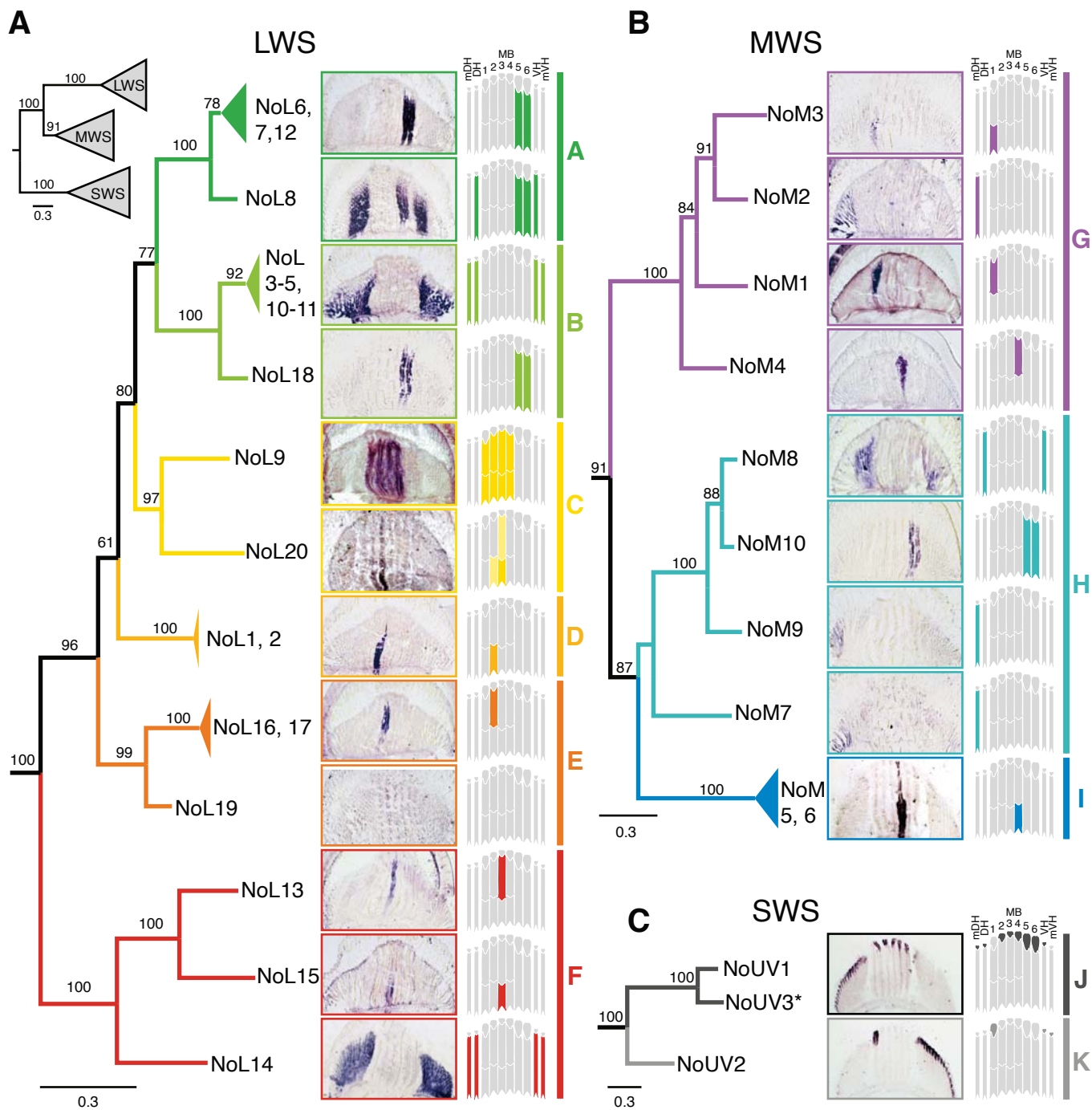
- 645 49. Awata H, Wakakuwa M, & Arikawa K (2009) Evolution of color vision in pierid butterflies:
646 blue opsin duplication, ommatidial heterogeneity and eye regionalization in *Colias erate*. *J*
647 *Comp Physiol A Neuroethol Sens Neural Behav Physiol* 195(4):401-408.
- 648 50. Shen WL, *et al.* (2011) Function of rhodopsin in temperature discrimination in *Drosophila*.
649 *Science* 331(6022):1333-1336.
- 650 51. Katana R, *et al.* (2019) Chromophore-Independent Roles of Opsin Apoproteins in *Drosophila*
651 Mechanoreceptors. *Curr Biol* 29(17):2961-2969 e2964.
- 652 52. Zanini D, *et al.* (2018) Proprioceptive Opsin Functions in *Drosophila* Larval Locomotion.
653 *Neuron* 98(1):67-74 e64.
- 654 53. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high
655 throughput. *Nucleic Acids Res* 32(5):1792-1797.
- 656 54. Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and
657 space complexity. *BMC Bioinformatics* 5:113.
- 658 55. Tamura K, Stecher G, Peterson D, Filipski A, & Kumar S (2013) MEGA6: Molecular
659 Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* 30(12):2725-2729.
- 660 56. Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with
661 thousands of taxa and mixed models. *Bioinformatics* 22(21):2688-2690.
- 662 57. Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of
663 large phylogenies. *Bioinformatics* 30(9):1312-1313.
- 664 58. Stamatakis A, Hoover P, & Rougemont J (2008) A rapid bootstrap algorithm for the RAxML
665 Web servers. *Syst Biol* 57(5):758-771.
- 666 59. Abascal F, Zardoya R, & Posada D (2005) ProtTest: selection of best-fit models of protein
667 evolution. *Bioinformatics* 21(9):2104-2105.
- 668 60. Darriba D, Taboada GL, Doallo R, & Posada D (2011) ProtTest 3: fast selection of best-fit
669 models of protein evolution. *Bioinformatics* 27(8):1164-1165.

- 670 61. Huelsenbeck JP & Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees.
671 *Bioinformatics* 17(8):754-755.
- 672 62. Ronquist F, *et al.* (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model
673 choice across a large model space. *Syst Biol* 61(3):539-542.
- 674 63. Miller M, Pfeiffer W, & Schwartz T (2010) Creating the CIPRES Science Gateway for
675 inference of large phylogenetic trees. Gateway Computing Environments Workshop (GCE),
676 2010. *IEEE*:1-8.
- 677 64. Miller MA, *et al.* (2015) A RESTful API for Access to Phylogenetic Tools via the CIPRES
678 Science Gateway. *Evol Bioinform Online* 11:43-48.
- 679 65. Guex N & Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for
680 comparative protein modeling. *Electrophoresis* 18(15):2714-2723.
- 681 66. Guex N, Peitsch MC, & Schwede T (2009) Automated comparative protein structure modeling
682 with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. *Electrophoresis* 30 Suppl
683 1:S162-173.
- 684 67. Johansson MU, Zoete V, Michielin O, & Guex N (2012) Defining and searching for structural
685 motifs using DeepView/Swiss-PdbViewer. *BMC Bioinformatics* 13:173.
- 686 68. Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol*
687 24(8):1586-1591.
- 688 69. Yang Z (1997) PAML: a program package for phylogenetic analysis by maximum likelihood.
689 *Comput Appl Biosci* 13(5):555-556.
- 690 70. Gao F, *et al.* (2019) EasyCodeML: A visual tool for analysis of selection using CodeML. *Ecol*
691 *Evol* 9(7):3891-3898.
- 692 71. Pond SL & Frost SD (2005) Datamonkey: rapid detection of selective pressure on individual
693 sites of codon alignments. *Bioinformatics* 21(10):2531-2533.

- 694 72. Poon AF, Frost SD, & Pond SL (2009) Detecting signatures of selection from DNA sequences
695 using Datamonkey. *Methods Mol Biol* 537:163-183.
- 696 73. Pond SLK & Muse SV (2005) HyPhy: hypothesis testing using phylogenies. *Statistical methods*
697 *in molecular evolution*, (Springer), pp 125-181.
- 698 74. Weaver S, *et al.* (2018) Datamonkey 2.0: a modern web application for characterizing selective
699 and other evolutionary processes. *Molecular biology and evolution* 35(3):773-777.
- 700 75. Cronin TW, Porter ML, Bok MJ, Wolf JB, & Robinson PR (2010) The molecular genetics and
701 evolution of colour and polarization vision in stomatopod crustaceans. *Ophthalmic and*
702 *Physiological Optics* 30:460-469.
- 703 76. Ishii T, Hirota J, & Mombaerts P (2003) Combinatorial coexpression of neural and immune
704 multigene families in mouse vomeronasal sensory neurons. *Curr Biol* 13(5):394-400.
- 705 77. Bodkin DK & Knudson DL (1985) Assessment of sequence relatedness of double-stranded
706 RNA genes by RNA-RNA blot hybridization. *J Virol Methods* 10(1):45-52.

A**B****D****C**





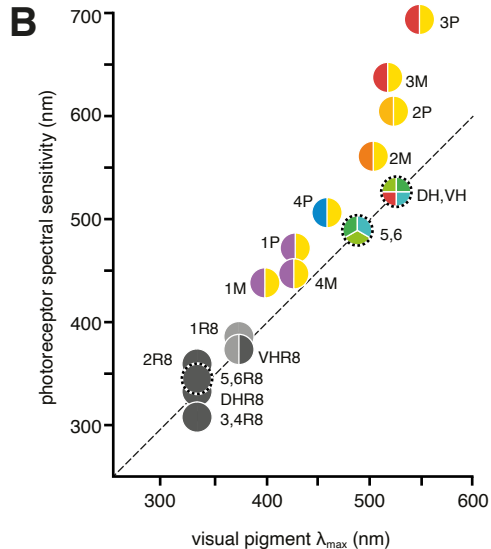
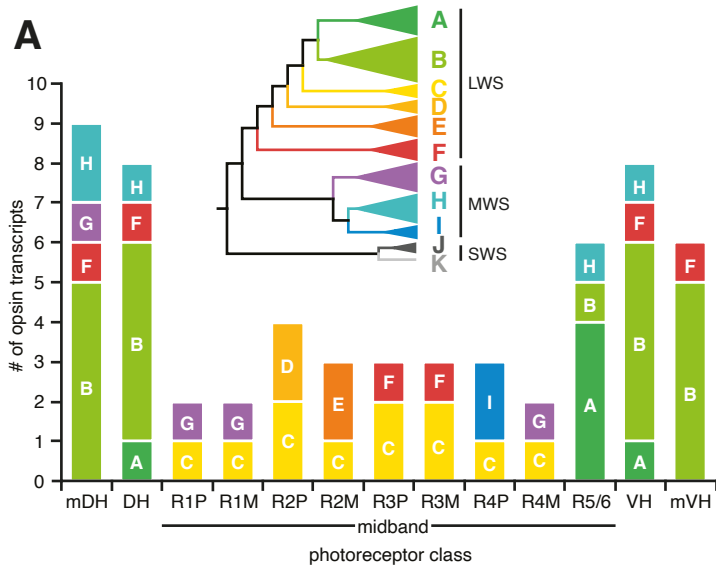


Table 1. Amino acid positions identified by proximity to the chromophore that contain S-A or F-Y substitutions (residues in bold). Amino acid sites are identified based on numbering in the NoL1 opsin of *Neogonodactylus oerstedii* (NoL1), the crystal structure of rhodopsin from the jumping spider *Hasarius adansoni* (JSR1; Accession # 6I9K_A), and bovine rhodopsin (BvRh; Accession # NP_001014890.1). Opsins are divided by spectral clades (long wavelength sensitive = LWS, middle wavelength sensitive = MWS, and ultraviolet sensitive = UVS) and subgroups (A-K). The number of sequences within each subgroup are indicated in the parentheses after each letter. Within each subgroup, all amino acids at that position are indicated, with the number of sequences encoding that amino acid indicated by subscript.

Amino Acid #s								
NoLI		110	129	134	194	205	226	318
JSR1		107	126	131	191	202	223	317
BVRh		94	113	118	178	189	210	292
LWS	A (4)	M ₃ L ₁	Y₄	N ₄	Y₄	T ₄	F₄	S₄
	B (6)	M ₆	Y₆	N ₆	Y₆	T ₆	F₆	S₅ ?₁
	C (2)	M ₂	Y₂	N ₂	Y₂	T ₂	F₂	S₂
	D (2)	L ₂	Y₂	N ₂	Y₂	T ₂	Y₂	S₂
	E (3)	M ₃	Y₃	S₃	Y₃	T ₃	Y₃	S₃
	F (3)	L ₂ V ₁	Y₃	S₃	Y₃	T ₃	F₁ Y₂	S₃
	G (2)	F₂	Y₂	G ₂	Y₂	F₂	Y₂	A₂
	H (4)	Y₃ F₁	Y₄	A₄	Y₄	F₃ Y₁	Y₄	A₄
	I (4)	Y₂ V ₁ L ₁	H ₂ F₁ Y₁	A₂ S₁ G ₁	F₃ Y₁	F₄	F₂ Y₂	G₃ A₁
	J (2)	F₂	F₁ Y₁	A₂	F₂	F₂	W ₂	A₂
UVS	K (1)	F₁	F₁	S₁	Y₁	F₁	Y₁	A₁

Table 2. Amino acid positions where residue variation corresponds to clade structure. The method of site identification is indicated: C = within 5Å of the chromophore; M = identified in MEME tests of selection, with a p<0.05. Amino acid sites are identified based on numbering in the NoL1 opsin of *Neogonodactylus oerstedii* (NoL1), the crystal structure of rhodopsin from the jumping spider *Hasarius adansoni* (JSR1; Accession # 6I9K_A), and bovine rhodopsin (BvRh; Accession # NP_001014890.1). Opsins are divided by spectral clades (long wavelength sensitive = LWS, middle wavelength sensitive = MWS, and ultraviolet sensitive = UVS) and subgroups (A-K). The number of sequences within each subgroup are indicated in the parentheses after each letter. Within each subgroup, all amino acids in that position are indicated, with the number of sequences indicated by subscript.

Amino Acid #s													
NoLI	106	118	138	204	218	221	222	225	250	296	299	321	
BVRh	103	115	135	201	215	218	219	222	247	294	297	320	
JSR1	90	102	122	188	202	204	206	209	236	269	271	295	
Method	C	M	C	C	C	C	C	C	M	C	C	C	
LWS	A (4)	N ₃ M ₁	T ₄	C ₄	G ₄	L ₄	Y ₄	T ₃ S ₁	V ₄	Q ₃ A ₁	L ₂ M ₂	N ₄	A ₄
	B (6)	M ₆	T ₆	C ₆	G ₆	L ₆	Y ₆	T ₆	V ₆	Q ₅ ? ₁	L ₅ ? ₁	N ₅ ? ₁	A ₅ ? ₁
	C (2)	M ₂	T ₂	C ₂	G ₂	L ₂	Y ₂	S ₂	V ₂	Q ₂	L ₂	N ₂	A ₂
	D (2)	M ₂	T ₂	C ₂	G ₂	L ₂	Y ₂	S ₂	V ₂	Q ₂	L ₂	N ₂	A ₂
	E (3)	M ₃	T ₃	C ₃	G ₃	L ₃	Y ₄	A ₂ S ₁	V ₃	Q ₃	F ₂ L ₁	N ₃	A ₄
	F (3)	M ₃	T ₃	C ₃	G ₃	L ₂ V ₁	Y ₃	T ₂ C ₁	C ₃	A ₁ N ₁ Q ₁	A ₁ C ₁ L ₁	N ₃	A ₃
MWS	G (2)	N ₂	V ₂	V ₂	S ₂	T ₂	I ₂	F ₂	D ₂	S ₂	A ₂	A ₂	A ₂
	H (4)	Q ₄	V ₄	L ₄	S ₄	G ₄	I ₃ L ₁	F ₄	D ₃ C ₁	A ₄	A ₄	T ₃ V ₁	A ₄
	I (4)	Q ₄	Y ₃ V ₁	D ₃ V ₁	S ₃ T ₁	G ₄	L ₂ M ₁ V ₁	F ₄	C ₃ Y ₁	A ₄	A ₄	V ₄	L ₄
UVS	J (2)	K ₂	G ₂	P ₂	T ₂	V ₂	L ₂	W ₂	M ₁ V ₁	A ₁ Q ₁	A ₂	A ₂	S ₁
	K (1)	K ₁	G ₁	I ₁	S ₁	V ₁	I ₁	F ₁	S ₁	A ₁	A ₁	A ₁	C ₂